Revisiting and Testing Stationarity in the Time-Frequency Plane

Patrick Flandrin[1]

Cnrs & École normale supérieure de Lyon
Physics Dept., Sisyph Group

[1] joint work with Pierre Borgnat, Jun Xiao, Paul Honeine, André Ferrari & Cédric Richard

Stationarity, from theory to practice

Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, . . .)

Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained

Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained
  1. stochastic framework, e.g., $\mathbb{E}x(t)\,x(t-\tau) = \gamma_x(\tau)$

Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained
  1. stochastic framework, e.g., $\mathbb{E}x(t)\,x(t-\tau) = \gamma_x(\tau)$
  2. invariance w.r.t. any time t and any shift τ

Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained
  1. stochastic framework, e.g., $\mathbb{E}x(t)\,x(t-\tau) = \gamma_x(\tau)$
  2. invariance w.r.t. any time t and any shift τ
- practice: definition loosely twisted

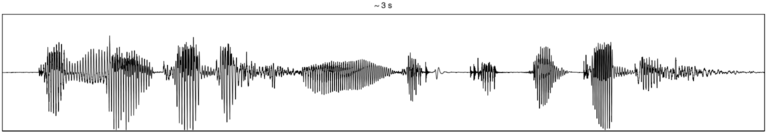
Stationarity, from theory to practice

- concept of stationarity
    - ubiquitous in signal/image processing
    - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained
    1. stochastic framework, e.g., $\mathbb{E}x(t)\,x(t-\tau) = \gamma_x(\tau)$
    2. invariance w.r.t. any time t and any shift τ
- practice: definition loosely twisted
    1. deterministic frameworks as well, e.g., $x(t) \sim x(t-kT)$

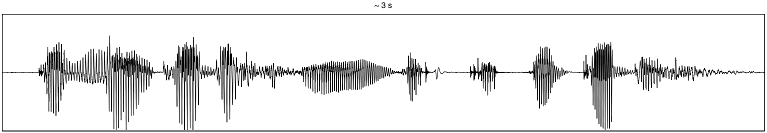
Stationarity, from theory to practice

- concept of stationarity
  - ubiquitous in signal/image processing
  - prerequisite for many tasks (analysis, modeling, ...)
- theory: definition heavily constrained
  1. stochastic framework, e.g., $\mathbb{E}x(t)\,x(t-\tau) = \gamma_x(\tau)$
  2. invariance w.r.t. any time t and any shift τ
- practice: definition loosely twisted
  1. deterministic frameworks as well, e.g., $x(t) \sim x(t-kT)$
  2. invariance w.r.t. selected times t and limited shifts (τ or kT)

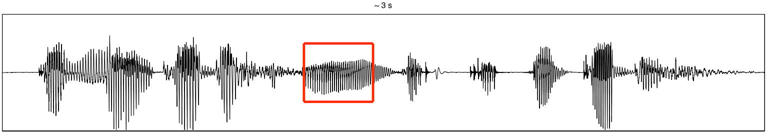
speech as an example



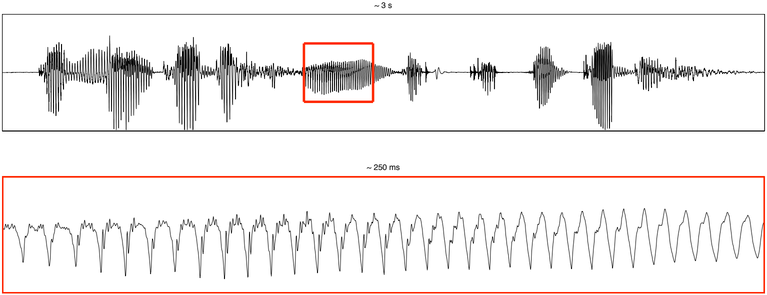
nonstationary



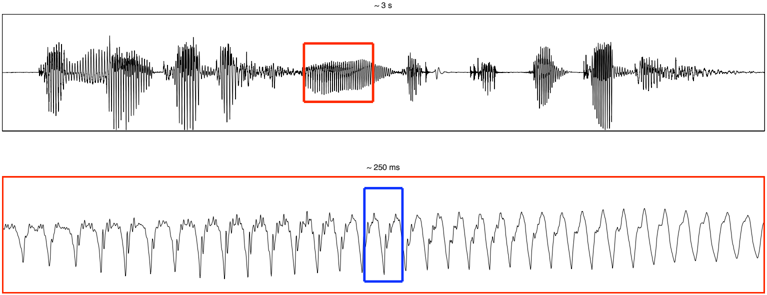
?


~ 3 s

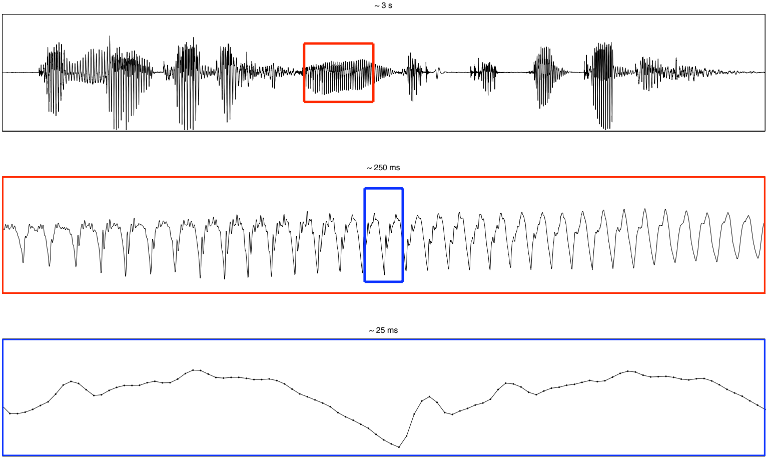
stationary



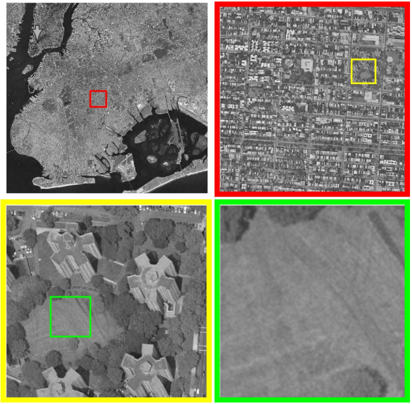
?


~ 3 s
~ 250 ms

nonstationary!



A 2D example



Objective and agenda

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates
3. distance-based test

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates
3. distance-based test
4. 3 variations

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates
3. distance-based test
4. 3 variations

   4.1 machine learning approach

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
  - encompassing both stochastic and deterministic situations
  - incorporating an observation scale in the definition
  - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates
3. distance-based test
4. 3 variations

   4.1 machine learning approach

   4.2 2D time-scale extension

Objective and agenda

- consider stationarity as a relative concept and develop an operational framework
    - encompassing both stochastic and deterministic situations
    - incorporating an observation scale in the definition
    - allowing for a possibility of a quantitative test

1. time-frequency framework
2. stationary reference from surrogates
3. distance-based test
4. 3 variations

    4.1 machine learning approach
    4.2 2D time-scale extension
    4.3 transient detection

A time-frequency approach 1.

A time-frequency approach 1.

- 2nd order stationarity: frequency description via the Power Spectrum Density (PSD)

$$\Gamma_x(f) := \int_{-\infty}^{+\infty} \gamma_x(\tau)\, e^{-i2\pi f\tau}\, d\tau$$

A time-frequency approach 1.

- 2nd order stationarity: frequency description via the Power Spectrum Density (PSD)

$$\Gamma_x(f) := \int_{-\infty}^{+\infty} \gamma_x(\tau)\, e^{-i2\pi f\tau}\, d\tau$$

- (harmonizable) nonstationary processes: PSD $\rightarrow$ Time-Varying Spectra (TVS) $\rho_x(t, f)$, with the key property: $\rho_x(t, f) = \Gamma_x(f), \forall t$ in the stationary case.

A time-frequency approach 2.

A time-frequency approach 2.

- estimation of TVS by means of multitaper spectrograms (or scalograms)

$$S_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K}\left|\int_{-\infty}^{+\infty} x(s)\,h_k(s-t)\,e^{-i2\pi fs}\,ds\right|^2,$$

$$\Omega_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K}\left|\int_{-\infty}^{+\infty} x(s)\,h_k\left(\frac{s-t}{f_0/f}\right)\,ds\right|^2,$$

with the K first Hermite (or Morse) functions used as short-time windows (or wavelets) $h_k(t)$

A time-frequency approach 2.

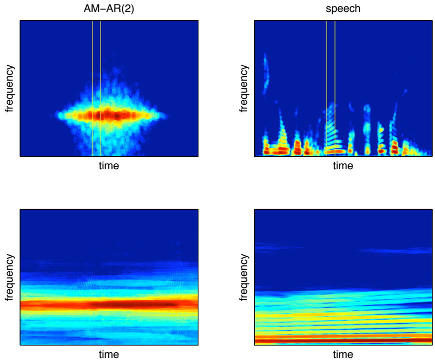
- estimation of TVS by means of multitaper spectrograms (or scalograms)

$$S_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K}\left|\int_{-\infty}^{+\infty} x(s)\,h_k(s-t)\,e^{-i2\pi fs}\,ds\right|^2,$$

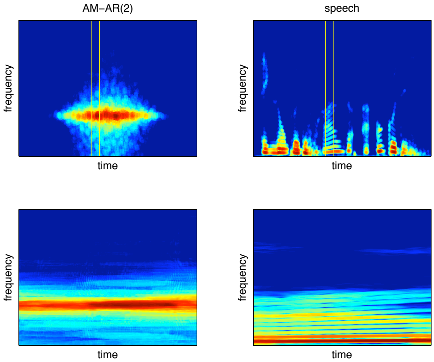
$$\Omega_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K}\left|\int_{-\infty}^{+\infty} x(s)\,h_k\left(\frac{s-t}{f_0/f}\right)\,ds\right|^2,$$

  with the K first Hermite (or Morse) functions used as short-time windows (or wavelets) $h_k(t)$
- rationale: "ensemble" averaging without any extra time averaging (conflicting with nonstationarity)

A time-frequency approach 3.

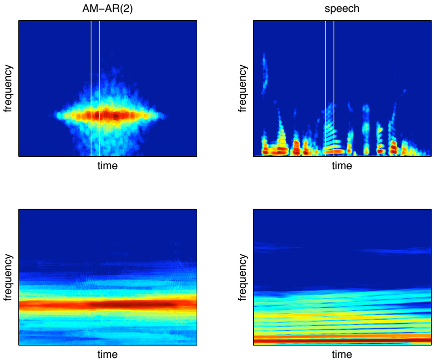


A time-frequency approach 3.



- advantage: common framework for both stochastic and deterministic situations

A time-frequency approach 3.



- advantage: common framework for both stochastic and deterministic situations
- rationale: relative stationarity = "homogeneity" within an observation scale $\Rightarrow$ comparison local vs. global

Local = global?

Local = global?

- test calls for a stationary reference: how?

Local = global?

- test calls for a stationary reference: how?
  - nonstationarity encoded in time evolution or, equivalently, in spectrum phase

Local = global?

- test calls for a stationary reference: how?
  - nonstationarity encoded in time evolution or, equivalently, in spectrum phase
  - stationarization via spectrum phase randomization (Richard *et al.*, IEEE-ICASSP'10)

Local = global?

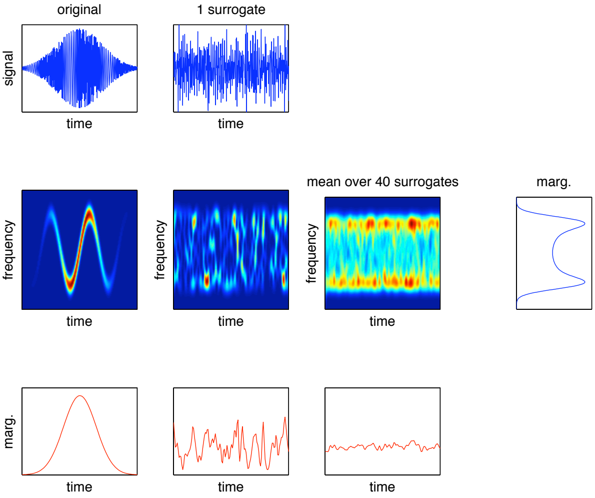
- test calls for a stationary reference: how?
  - nonstationarity encoded in time evolution or, equivalently, in spectrum phase
  - stationarization via spectrum phase randomization (Richard *et al.*, IEEE-ICASSP'10)
- new use of surrogate data technique (Theiler *et al.*, '92)

Local = global?

- test calls for a stationary reference: how?
  - nonstationarity encoded in time evolution or, equivalently, in spectrum phase
  - stationarization via spectrum phase randomization (Richard *et al.*, IEEE-ICASSP'10)
- new use of surrogate data technique (Theiler *et al.*, '92)
- basic algorithm:

$\hat{x} = \mathrm{FFT}(x)$ % x = original data
draw WGN $\epsilon(t)$ and compute $\hat{\epsilon} = \mathrm{FFT}(\epsilon)$
$\hat{x} \leftarrow |\hat{x}| \exp\{i \arg \hat{\epsilon}\}$
$y = \mathrm{IFFT}(\hat{x})$ % y = surrogate data

Stationarization via surrogates



The proposed approach

[Xiao *et al.*, EUSIPCO'07] [Xiao *et al.*, IEEE-SSP'07]
[Borgnat *et al.*, IEEE T-SP'10]

The proposed approach

[Xiao *et al.*, EUSIPCO'07] [Xiao *et al.*, IEEE-SSP'07]
[Borgnat *et al.*, IEEE T-SP'10]

1. compute, from the data, a set of stationary surrogates

The proposed approach

[Xiao *et al.*, EUSIPCO'07] [Xiao *et al.*, IEEE-SSP'07]
[Borgnat *et al.*, IEEE T-SP'10]

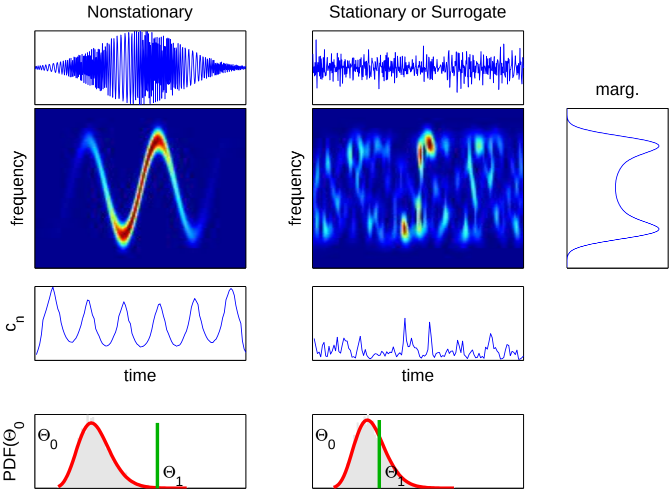
1. compute, from the data, a set of stationary surrogates
2. attach to both data and surrogates a series of features aimed at comparing local vs. global behaviors

The proposed approach

[Xiao *et al.*, EUSIPCO'07] [Xiao *et al.*, IEEE-SSP'07]
[Borgnat *et al.*, IEEE T-SP'10]

1. compute, from the data, a set of stationary surrogates
2. attach to both data and surrogates a series of features aimed at comparing local vs. global behaviors
3. construct a test based on the empirical statistical characterization of such features for surrogates (null hypothesis of stationarity)

Principle of distance-based test



More on distance-based test 1.

- comparison local vs. global for the data

$$c_n^{(x)} := D\left(S_{x,K}(t_n,\cdot), \langle S_{x,K}(t_n,\cdot)\rangle_n\right),$$

  with $D(\cdot,\cdot)$ some dissimilarity measure
- creation of a set of surrogates and similar comparisons

$$x(t) \to \{s_j(t), j = 1, \ldots J\}$$

$$\{c_n^{(s_j)} := D\left(S_{s_j,K}(t_n,\cdot), \langle S_{s_j,K}(t_n,\cdot)\rangle_n\right), j = 1, \ldots J\}$$

More on distance-based test 2.

- measure the ℓ_2 fluctuations of D for data and surrogates

$$\Theta_1 = L\left(c^{(x)}, \langle c^{(x)}\rangle_n\right); \left\{\Theta_0(j) = L\left(c^{(s_j)}, \langle c^{(s_j)}\rangle_n\right), j = 1, \ldots J\right\}$$

with

$$L(g,h) := \frac{1}{N}\sum_{n=1}^{N}(g_n - h_n)^2$$

- construct the one-sided test

$$\begin{cases} \Theta_1 > \gamma & : \text{“\textit{nonstationarity}”} \\ \Theta_1 < \gamma & : \text{“\textit{stationarity}”} \end{cases}$$

with γ some threshold derived from the empirical pdf of Θ_0

Associated quantities

- index of nonstationarity

$$\mathrm{INS} := \sqrt{\frac{\Theta_1}{\langle \Theta_0(j) \rangle_j}}$$

- scale of nonstationarity

$$\mathrm{SNS} := \frac{1}{T} \arg\max_{T_h} \{\mathrm{INS}(T_h)\},$$

with T the observation span and T_h the window length

Choosing a distance

- typical nonstationarities captured by time-varying spectra: AM (level change) and FM (shape change)
- motivates a combination of log-spectral deviation (AM) and Kullback-Leibler divergence (FM)

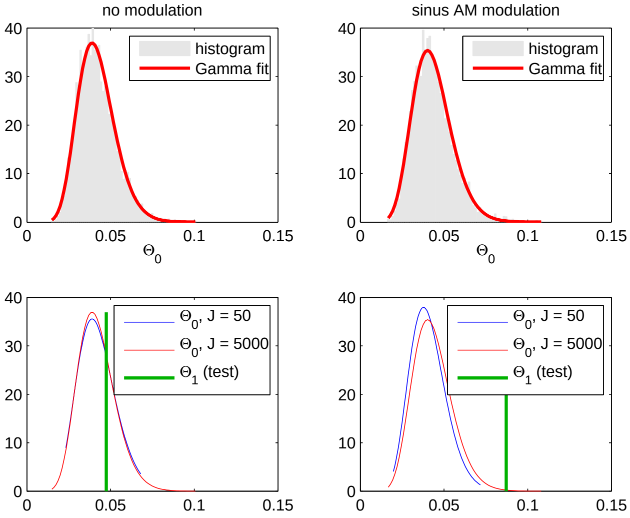
$$D(G,H) := D_{\mathrm{KL}}(\tilde{G},\tilde{H}).\,(1 + D_{\mathrm{LSD}}(G,H))\,,$$

with

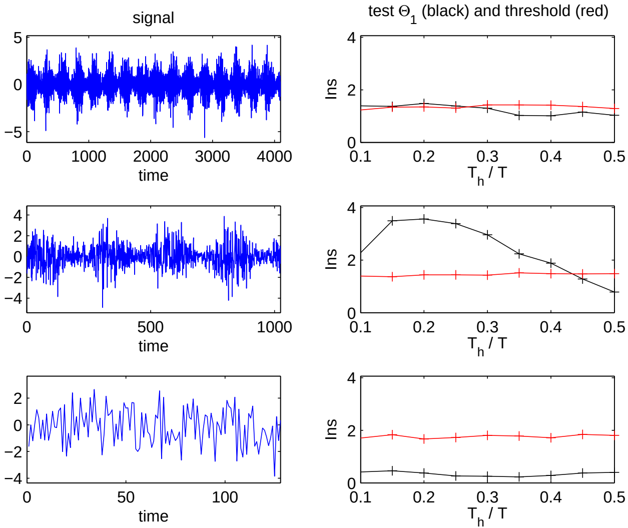
$$D_{\mathrm{KL}}(G,H) := \int_{\Omega} (G(f) - H(f)) \log \frac{G(f)}{H(f)}\, df$$

$$D_{\mathrm{LSD}}(G,H) := \int_{\Omega} \left| \log \frac{G(f)}{H(f)} \right| df$$

Choosing surrogates



Synthetic data



Variation 1: machine learning

- approaches via distances are model-based since they require some (parametric) knowledge of surrogates features pdf
- possible way out by considering surrogates as a learning set attached to stationarity
- stationarity test recast as outlier detection by using the machinery of one-class SVM

Principle of SVM-based test

- rationale: determine the minimum volume hypersphere that encloses (most of) the training points, up to a small fraction of data excluded from the domain.
- optimization: trade-off between minimizing the radius r^* of the enclosing hypersphere and controlling the sum of the slack variables ξ_i^* associated with each outlier.

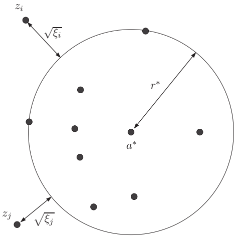
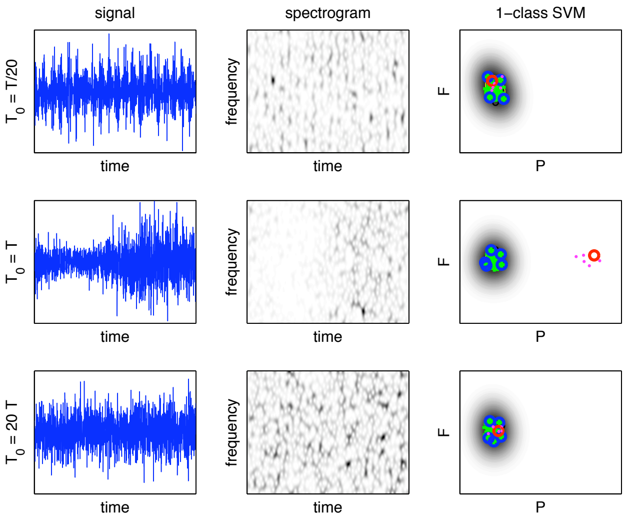


Illustration of SVM-based test



Variation 2: 2D extensions

Variation 2: 2D extensions

- extension 1: apply the same strategies to 2D TF spectra

Variation 2: 2D extensions

- extension 1: apply the same strategies to 2D TF spectra
- extension 2: replace *mutatis mutandis* TF by space-scale (e.g., spectrograms $\rightarrow$ scalograms)

Variation 2: 2D extensions

- extension 1: apply the same strategies to 2D TF spectra
- extension 2: replace *mutatis mutandis* TF by space-scale (e.g., spectrograms $\rightarrow$ scalograms)
    1. multiresolution $\Rightarrow$ selection of observation scale

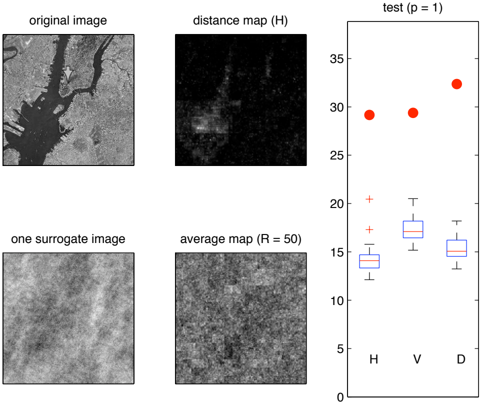
Variation 2: 2D extensions

- extension 1: apply the same strategies to 2D TF spectra
- extension 2: replace *mutatis mutandis* TF by space-scale (e.g., spectrograms $\rightarrow$ scalograms)
  1. multiresolution $\Rightarrow$ selection of observation scale
  2. possibility of directional tests

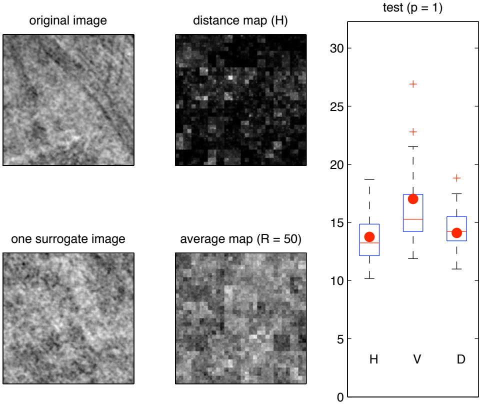
Variation 2: 2D extensions

- extension 1: apply the same strategies to 2D TF spectra
- extension 2: replace *mutatis mutandis* TF by space-scale (e.g., spectrograms $\rightarrow$ scalograms)
  1. multiresolution $\Rightarrow$ selection of observation scale
  2. possibility of directional tests
  3. here: undecimated dyadic ("symmlet-4") tensor wavelet transform, with test based on the ℓ_1-norm of the mixed distance map (Kullback-Leibler + log-spectral deviation) computed pointwise in the 3 directions

Back to the 2D example



Back to the 2D example



Variation 3: transient detection

Variation 3: transient detection

1. TF model = localized events in smoothly spread noise

Variation 3: transient detection

1. TF model = localized events in smoothly spread noise
2. In practice, only one observation
   $\Rightarrow$ statistical fluctuations in the estimated noise background
   $\Rightarrow$ false transients

Variation 3: transient detection

1. TF model = localized events in smoothly spread noise
2. In practice, only one observation
   $\Rightarrow$ statistical fluctuations in the estimated noise background
   $\Rightarrow$ false transients
3. Way out = compare data to a TF stationarized reference
   $\Rightarrow$ surrogates from a 2D phase randomization with a positivity constraint (spectrogram)

Variation 3: transient detection

1. TF model = localized events in smoothly spread noise
2. In practice, only one observation
   $\Rightarrow$ statistical fluctuations in the estimated noise background
   $\Rightarrow$ false transients
3. Way out = compare data to a TF stationarized reference
   $\Rightarrow$ surrogates from a 2D phase randomization with a positivity constraint (spectrogram)
4. Detection via an entropy measure (Rényi)

Algorithm

$A_x = \text{2D-FFT}(S_x)$ % S_x = spectrogram

draw WGN $\epsilon(t)$ and compute $A_\epsilon = \text{2D-FFT}(S_\epsilon)$

$A_x \leftarrow |A_x| \exp\{i \arg A_\epsilon\}$ % A_x = surrogate ambiguity function

test = test$_0 >$ thresh

$r = 0$

while test $\geq$ thresh **do**

  $r \leftarrow r + 1$

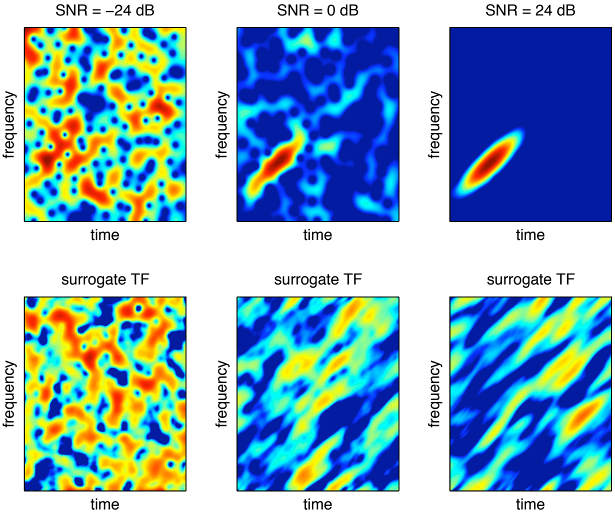
  draw WGN $\epsilon(t)$ and compute $A_\epsilon = \text{2D-FFT}(S_\epsilon)$

  $A_x = \text{2D-FFT}([\text{2D-IFFT}(A_x)]_+)$

  $A_x \leftarrow |A_x| \exp\{i(\arg A_x + \lambda^r \arg A_\epsilon)\}$

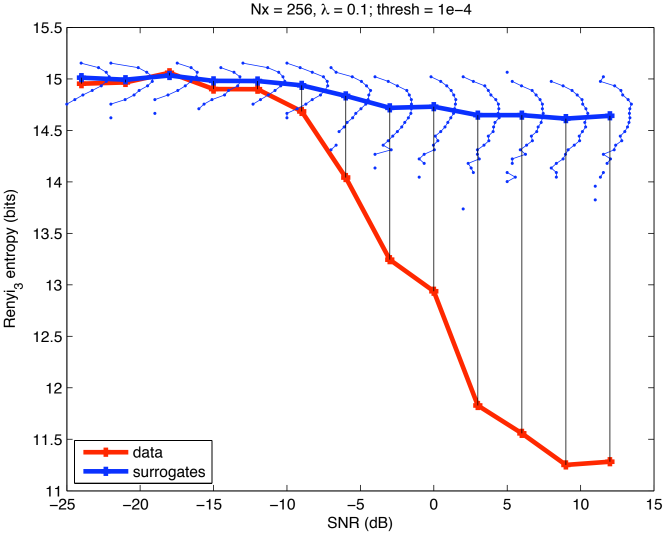
  test $\leftarrow \text{vol}(S_x < 0)/\text{vol}(S_x)$

Example



SNR = −24 dB
SNR = 0 dB
SNR = 24 dB
surrogate TF
surrogate TF
surrogate TF
frequency
time

Performance



Concluding remarks

Concluding remarks

- stationarity revisited from an operational perspective

Concluding remarks

- stationarity revisited from an operational perspective
  1. explicitly considered as a relative concept
  2. tested based on data-driven surrogate data

Concluding remarks

- stationarity revisited from an operational perspective
  1. explicitly considered as a relative concept
  2. tested based on data-driven surrogate data
- further variations
  1. from detection to classification [Amoud *et al.*, GRETSI'09]
  2. global, features-free, learning strategies [Amoud *et al.*, IEEE-SSP'09]
  3. softened transitional surrogates [Borgnat *et al.*, ICASSP'11]

Concluding remarks

- stationarity revisited from an operational perspective
  1. explicitly considered as a relative concept
  2. tested based on data-driven surrogate data
- further variations
  1. from detection to classification [Amoud *et al.*, GRETSI'09]
  2. global, features-free, learning strategies [Amoud *et al.*, IEEE-SSP'09]
  3. softened transitional surrogates [Borgnat *et al.*, ICASSP'11]
- extension to generalized forms of stationarity (Lamperti)

More

(p)reprints and Matlab codes available at

`http://perso.ens-lyon.fr/patrick.flandrin`